

Análisis de Dispositivos

# Dispositivos de Interacción Mediante Reconocimiento de Voz

## Índice

Nombre genérico y comercial.....	Pág.3
Tipo de Interacción.....	Pág.3
Paradigma de Interacción.....	Pág.4
Estado comercial.....	Pág.5
Ventajas de su Uso.....	Pág.6
Posibles Desventajas.....	Pág.7
Integración con otros dispositivos.....	Pág.7
Posibles usos del dispositivo.....	Pág.8-9

## Nombre genérico y comercial

Este tipo de dispositivos se denominan por dispositivos de Reconocimiento Automático del Habla (RAH) o Reconocimiento Automático de voz.

En cuanto a nombres comerciales hay una gran variedad dado que existen múltiples softwares orientados a esta aplicación, entre los que podemos destacar tales como IBM Embedded ViaVoice, Dragon, Siri, el propio Windows Vista Reconocimiento de Voz, el Voice Tools Project de Eclipse y los creados por la empresa Phonetic Arts para dispositivos Android.

## Tipo de Interacción

El tipo de interacción principal de estos dispositivos es el de entrada-salida.

El usuario interactúa con los dispositivos, principalmente, mediante su voz como único elemento para comunicarse, por lo que el habla sería la única entrada hacia el dispositivo. Como respuesta a la entrada recibida, el dispositivo responde mediante alguna salida, tal como escritura, accesos a otras aplicaciones, etc.

Ejemplo: En el caso del dictado por voz, el usuario dicta mediante el habla el contenido a escribir (entrada), y como consecuencia el ordenador escribe automáticamente el texto dictado (salida) mostrándolo por pantalla.

## Paradigma de Interacción

Generalmente estos dispositivos son o bien ordenadores o dispositivos móviles, por lo que la interacción sería directa entre usuario-computador. El modo de uso sería, en algunos casos, a través de la activación de la aplicación, mediante un botón, y tras ella, la comunicación del usuario mediante el habla. Esto generalmente es para los casos en los que la aplicación se usa en casos puntuales; por ejemplo para usar la aplicación de marcación por voz en un teléfono móvil, el usuario ha de mantener un botón pulsado hasta que recibe un aviso, por audio o vibración, que indica que la aplicación está activada y entonces dice el nombre del contacto a llamar, el móvil lo reconoce y ejecuta la llamada.

En otros casos en los que la aplicación se usa de forma habitual, no es necesario ejecutar la aplicación cada vez que se requiere realizar una acción, sino que esta está en constante recepción de forma que cada vez que el usuario habla, éste lo recibe y ejecuta la acción. Este caso sería por ejemplo el reconocimiento de voz de Windows, que se activa una vez iniciado el sistema y el usuario va ejecutando comandos cada vez que quiere sin necesidad de estar pendiente de 'avisar' al sistema de que va a hablar.

## Estado comercial

Estos dispositivos, se empezaron a desarrollar sobre los años 50 y empezaron a comercializarse sobre 1990 en el ámbito de las computadoras, pero, aunque parezca un gran avance, la mayoría de las personas no son usuarios habituales dado que prefieren, o están habituados, a utilizar el ratón y teclado. Aunque ésto les pueda parecer más cómodo y rápido el hecho de dictar mediante voz sería mucho más rápido que teclearlo.

En cambio, en los sistemas portátiles quizás sea donde los usuarios hagan más uso de este tipo de dispositivos. Desde hace años ya se disponía de aplicaciones básicas tales como la marcación por voz, que junto al uso del manos libres nos permitían hacer llamadas sin utilizar las manos en absoluto. Actualmente la industria de la telefonía móvil está en constante desarrollo lo que supone el lanzamiento de nuevas aplicaciones que han ido aumentando las posibles acciones a realizar mediante el uso de voz. Cabe destacar el reciente lanzamiento de Siri, aplicación desarrollada por Apple, que proporciona la ayuda de un 'asistente' el cual, mediante la voz, te ayuda a realizar cualquier acción con tu dispositivo.

## Ventajas de su Uso

Como ventaja principal, y posiblemente la más significativa, es la de que permite el acceso a cierto tipo de aplicaciones a las personas con discapacidad sensorial o cognitiva las cuales, sin estos dispositivos, no podían acceder:

Permite la utilización de dispositivos a personas con deficiencias visuales, motoras y cognitivas, ya que evita el uso de tener que escribir con teclados, deletrear, pulsar teclas, etc.

Como ventajas generales:

- Agiliza la escritura de textos, dado que es más rápido hablar que teclear.
- Permite realizar las tareas con mas comodidad al tener las manos y la vista despejadas.
- No requiere sistemas de aprendizaje dado que el dialogo es algo natural para los usuarios.
- En dispositivos portátiles permite la realización de tareas durante desplazamientos. El usuario puede andar o correr, sin atender a una pantalla o teclado.
- Ahorro de batería en dispositivos dado que no necesitan la utilización de pantallas durante su uso.

## Posibles Desventajas

- Posibles fallos y confusión de palabras.
- Tiempo de 'entrenamiento'.
- Problemas en ambientes ruidosos.
- Problemas con la velocidad del habla.
- Carencia de vocabulario al ser éste tan extenso o con modismos que puedan crear confusiones.
- Falta de privacidad.
- En algunos casos pérdida de información.

## Integración con otros dispositivos

Los dispositivos de reconocimiento de voz pueden integrarse con una gran variedad de dispositivos.

Como dispositivo básico para usar un software de éstas características necesitaríamos un dispositivo de entrada de audio, un micrófono y para poder recibir una respuesta también acústica una salida de audio, unos altavoces. Habitualmente estos dispositivos están integrados en computadores, dispositivos móviles de telefonía, relojes, etc. por lo que su uso generalmente suele estar complementado con teclados y pantallas.

Podrían integrarse con cualquier otro dispositivo siempre que se pudieran sustituir comandos de teclas por otros de voz.

## Posibles usos del dispositivo

-*Manos libres*: en uno de los usos más generalizados dado que permite al usuario realizar tareas sin tener que emplear las manos ni prestar atención visual.

-*Dictado*: este uso complementa a los editores de texto dando la posibilidad de ir dictando el texto a escribir sin tener que teclearlo, lo que lo hace más rápido y menos pesado, en especial cuando los textos son largos.

-*Control de comandos*: incluidos en un SO permiten la posibilidad de ejecutar comandos mediante instrucciones sencillas (del tipo 'abrir+programa', 'cerrar+programa'). Esto agiliza la forma de ejecutarlos dado que apenas dos palabras sustituirían la acción de abrir menús y hacer varios clics con el ratón .

-*Telefonía*: hoy en día en la mayoría de teléfonos de atención al cliente en los que atiende directamente un operador automático la forma de seleccionar los menús suele ser pulsando números, lo que conlleva, en los teléfonos móviles, tener que despegar el teléfono de la oreja y apretar la tecla cada vez que sale una opción, o tener que desviar la mirada al teclado del teléfono fijo para marcarlo. Con la opción de reconocimiento de voz, en vez de pulsar el número, puedes decirlo, o en algunos servicios, decir directamente qué es lo que quieres hacer evitando todos los menús previos.



*-Personas con discapacidad:* Este uso creo que debería ser uno de los más valorados debido a que permite acceder a ciertos servicios, a personas las cuales antes lo veían como algo totalmente inviable. De esta forma, una persona ciega puede acceder a un dispositivo (ordenador, teléfono móvil...) sin necesidad de que éste esté adaptado físicamente a ello (no necesitaría teclados en Braille). También podría recibir SMS mediante síntesis de voz. Las personas con problemas motores se evitarían de esta forma tener que pulsar botones, que en algunos casos son diminutos y poco accesibles. Y por último, para las personas con deficiencias auditivas podrían utilizar este tipo de dispositivos para recibir llamadas y poder obtener texto escrito del habla.